<http://www.diva-portal.org>

Preprint

This is the submitted version of a paper published in *Carbohydrate Research*.

Citation for the original published paper (version of record):

Shashkov, A S., Zhang, W., Perepelov, A V., Weintraub, A., Liu, B. et al. (2016)
Structure of the O-polysaccharide of *Escherichia coli* O132
Carbohydrate Research, 427: 44-47
<https://doi.org/10.1016/j.carres.2016.03.016>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-131489>

Note

Structure of the O-polysaccharide of *Escherichia coli* O132

Alexander S. Shashkov,^a Wenwen Zhang,^b Andrei V. Perepelov,^{a,*} Andrej Weintraub,^c Bin Liu,^b
Göran Widmalm,^d Yuriy A. Knirel^a

^a*N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 119991 Moscow,
Russian Federation*

^b*TEDA Institute of Biological Sciences and Biotechnology, Nankai University, TEDA,
300457 Tianjin, China*

^c*Department of Laboratory Medicine, Division of Clinical Microbiology,
Karolinska Institute, Karolinska University Hospital, S-141 86 Stockholm, Sweden*

^d*Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University,
S-106 91 Stockholm, Sweden*

Abstract

Mild acid degradation of the lipopolysaccharide of *Escherichia coli* O132 released its O-polysaccharide. Analysis by 1D and 2D ¹H and ¹³C NMR spectroscopy prior and subsequent to O-deacetylation, in conjunction with sugar analysis, revealed a linear pentasaccharide repeating unit of the O-polysaccharide having the following structure:

Putative functions of genes in the O-antigen gene cluster of *E. coli* O132 are consistent with the O-polysaccharide structure.

Keywords: *Escherichia coli*; lipopolysaccharide; O-polysaccharide; bacterial polysaccharide structure; O-antigen gene cluster

Escherichia coli is a gram-negative bacterium that colonizes the gastrointestinal tract of humans and warm-blooded animals where it usually is a harmless commensal beneficial to the host. However, there are pathogenic variants of *E. coli* clones that have acquired virulence factors thereby increasing

their ability to adapt to new niches.^{1,2} These pathogenic *E. coli* cause diseases such as diarrhea, septicemia and urinary tract infection. *E. coli* O132 is a serogroup that has been reported to cause septicemia in chickens,³ diarrhea in rabbits where its pathotype was enteropathogenic *E. coli*,⁴ diarrhea in calves where the pathotype was Shiga toxin-producing *E. coli*,⁵ and diarrhea in children under three years of age (enterotoxigenic *E. coli*).⁶ We herein elucidate the structure of the *E. coli* O132 O-antigen polysaccharide using chemical, NMR spectroscopic and bioinformatics methods.

Lipopolysaccharide (LPS) was isolated from bacterial cells of *E. coli* O132 and hydrolyzed with mild acid to give an O-polysaccharide, which was isolated by GPC. Sugar analysis by GLC of the alditol acetates derived after full acid hydrolysis of the polysaccharide revealed Rha, Glc, Gal, and GlcNAc in the ratios 1.9:1:0.8:0.6 (detector response). The L configuration of Rha and the D configuration of the other monosaccharides were not determined chemically but inferred from the ¹³C NMR chemical shifts⁷ and the content of the O-antigen gene cluster (see below).

The ¹³C NMR spectrum of the O-polysaccharide (Fig. 1, bottom) showed signals for five anomeric carbons in the region δ 98.4-102.9, three C-CH₂OH groups (C-6 of Glc, Gal, and GlcN) at δ 61.0-64.0, two C-CH₃ groups (C-6 of Rha) at δ 17.7-17.8, one nitrogen-bearing carbon (C-2 of GlcN) at δ 56.7, 19 oxygen-bearing sugar ring carbons in the region δ 70.2-84.7, one *N*-acetyl group at δ 23.6 (Me) and 175.7 (CO), and one *O*-acetyl group at δ 21.9 (Me) and 174.5 (CO). The ¹H NMR spectrum showed signals for five anomeric protons at δ 4.58-5.18, two C-CH₃ groups (H-6 of Rha) at δ 1.26-1.28, other sugar protons in the region δ 3.34-4.35, one *N*-acetyl group at δ 2.07, and one *O*-acetyl group at δ 2.24. Therefore, the O-antigen has a pentasaccharide repeating unit containing one residue each of Glc, Gal, and GlcNAc, two residues of Rha, and one *O*-acetyl group.

The O-polysaccharide was O-deacetylated with aq ammonia, and the ¹H and ¹³C NMR spectra of the native and O-deacetylated O-polysaccharides were assigned using 2D ¹H,¹H-COSY, TOCSY, ¹H,¹³C-HSQC, and HSQC-TOCSY experiments (Table 1). Based on ¹³C NMR chemical shifts, intra-residue ¹H,¹H correlations, and ³J_{H,H} coupling constant values estimated from the 2D NMR spectra, five spin-systems were identified, including one each for α -Glc_p (unit **C**), α -Gal_f (unit **E**), and β -Glc_pNAc (unit **A**), and two for α -Rha_p (units **B** and **D**). The spin system for GlcNAc was recognized by correlations between the proton at the nitrogen-bearing carbon (H-2) at δ 3.74 and the corresponding carbon (C-2) at δ 56.7 in the ¹H,¹³C-HSQC spectrum. A relatively large J_{1,2} coupling constant ~8 Hz confirmed that unit **A** is β -linked. The C-1 chemical shift δ 102.3 of unit **E** confirmed the α -linkage of the Gal_f residue, and the C-5

chemical shifts δ 72.3 of unit **C** and δ 70.2-70.5 of units **B** and **D** corroborated the α -linkages of the three other monosaccharides⁸ in the native O-polysaccharide.

The 2D ^1H , ^1H -ROESY spectrum of the O-deacetylated polysaccharide (Fig. 2) showed **A** H-1/**E** H-2, **B** H-1/**A** H-3, **D** H-1/**C** H-4, and **E** H-1/**D** H-3 cross-peaks between the transglycosidic protons. The 2D ^1H , ^{13}C -HMBC experiment (Fig. 3) revealed the following correlations between the anomeric protons and the linkage carbons: **A** H-1/**E** C-2, **B** H-1/**A** C-3, **D** H-1/**C** C-4, and **E** H-1/**D** C-3. These data were in agreement with downfield displacements, due to glycosylation,⁹ of the signals for the linkage carbons of units **A–E** (Table 1) as compared with their positions in the corresponding non-substituted monosaccharides.⁸ Therefore, the O-deacetylated polysaccharide has the following structure:

A comparison of the ^1H , ^{13}C -HSQC spectra of the native and O-deacetylated O-polysaccharides showed a downfield displacement, due to a deshielding effect, of the H-2/C-2 cross-peak of Rha **D** from δ 4.15/69.2 in the latter to δ 5.08/72.4 in the former. Accordingly, in the ^{13}C NMR spectrum of the native O-polysaccharide (Fig. 1, bottom), the signal for C-1 of unit **D** shifted upfield to δ 98.4 as compared with its position at δ 101.6 in the spectrum of the O-deacetylated polysaccharide (Fig. 1, top) (β -effect of O-acetylation^{10,11}). Therefore, Rha **D** is 2-O-acetylated, and the O-polysaccharide of *E. coli* O132 has the following structure:

The O-antigen gene cluster of *E. coli* O132 between conserved genes *galF* and *gnd* (GenBank accession number AB812056)¹² includes 12 genes (*orf1-12*), which have the same transcriptional direction from *galF* to *gnd*. D-Glc, D-Gal, and D-GlcNAc are common sugars in bacteria, and genes for synthesis of their nucleotide precursors are usually located outside the O-antigen gene cluster. However, *glf* gene (*orf6*) for conversion of UDP-D-Galp into UDP-D-Galf is present and identified by high level homology to *glf* genes of other bacteria.¹³ At the 5'-end of the O-antigen gene cluster, there are four *rml* genes (*orf1-4*) typically grouped together in the order *rmlBDAC*,¹⁴⁻¹⁶ which are responsible for the synthesis of dTDP-L-Rha.

The presence of *wzx* gene (*orf5*) for flippase responsible for translocation of the lipid-linked O-unit through the inner membrane, and *wzy* gene (*orf10*) for O-antigen polymerase that mediates O-unit polymerization indicates that the O132-antigen is synthesized by the Wzx/Wzy-dependent pathway.¹⁷ There are four putative glycosyltransferase genes (*orf7,8,11,12*) for the O-

unit assembly. One from the predicted glycosyltransferases (*orf7*) is 43% identical and 64% similar to rhamnosyl transferase Orf11 (WbbL) of *E.coli* O16 (GenPept accession number BAQ00657.1), which catalyzes formation of the α -L-Rhap-(1 $\rightarrow$ 3)-D-GlcpNAc linkage.¹⁸ This GT is also similar to WekH in *E. coli* O119 and WfaL in *E. coli* O139, both proposed to carry out the transfer of L-Rhap to O3 of D-GlcpNAc thereby forming the disaccharide structural element α -L-Rhap-(1 $\rightarrow$ 3)-D-GlcpNAc,¹⁹ present in the *E. coli* O119 and O139 O-antigens.^{20,21} The other glycosyltransferases were not assigned to the linkages but their total number is in accord with the pentasaccharide O-unit. Finally, one acetyltransferase gene (*orf9*) is present. Therefore, the O-polysaccharide structure is consistent with putative functions of genes in the O-antigen gene cluster of *E. coli* O132.

1. Experimental

1.1. Bacterial strain and isolation of the lipopolysaccharide

E. coli O132:K+:H28 strain CCUG 11432 was obtained from the Culture Collection, University of Göteborg, Sweden. The bacteria were grown and the LPS was isolated essentially as described.²²

1.2. Degradation of the lipopolysaccharide

Mild acid hydrolysis of the LPS (117 mg) was performed with aq 2% AcOH (2 mL) at 100 °C until precipitation of a lipid (1 h). A polysaccharide (46 mg) was isolated from the supernatant by gel permeation chromatography (GPC) on a column (56 $\times$ 2.6 cm) of Sephadex G-50 Superfine (Amersham Biosciences, Sweden) in 0.05 M pyridinium acetate buffer pH 4.5 monitored by a differential refractometer (Knauer, Germany).

An O-polysaccharide sample (45 mg) was treated with aq 25% ammonia (2 mL) at 37 °C for 8 h and an O-deacylated polysaccharide (23 mg) was isolated from the supernatant by GPC on Sephadex G-50 Superfine as described above.

1.3. Chemical analyses

A sample of the O-polysaccharide (0.5 mg) was hydrolyzed with 2 M CF₃CO₂H (120 °C, 2 h). Monosaccharides were conventionally converted into the alditol acetates and identified by GLC on a Maestro 7820 GC instrument (Interlab, Russia) equipped with a HP-5ms column using a temperature program of 160 °C (1 min) to 290 °C at 7 °C min⁻¹.

1.4. NMR spectroscopy

Samples were deuterium-exchanged by freeze-drying from 99.9 % D₂O and then examined as solutions in 99.95% D₂O. NMR spectra were recorded on a Bruker Avance II 600 MHz spectrometer (Germany) at 30 °C. ¹H and ¹³C NMR chemical shifts were measured with internal sodium 3-(trimethylsilyl)propanoate-2,2,3,3-d₄ (δ_{H} 0, δ_{C} -1.6) as reference. 2D NMR spectra were obtained using standard Bruker software, and the Bruker TopSpin 2.1 program was used to acquire and process the NMR data. A spin-lock time of 60 ms and a mixing time of 150 ms were used in the 2D TOCSY and ROESY experiments, respectively. The ¹H, ¹³C-HMBC experiment was performed with a 60-ms delay for the evolution of long-range connectivities.

Acknowledgements

This work was supported by the Russian Foundation for Basic Research (project 14-04-01709), the National 973 Project of China (Nos. 2012CB721101 and 2013CB733904), the International Science & Technology Cooperation Program of China (2013DFR30640), National Natural Science Foundation of China (NSFC) Program (31530083, 31270003, 81471904, 31371259, 31270133 and 31470194), and Research Project of Chinese Ministry of Education (NO.113015A), and the Swedish Research Council.

References

1. Stenutz R, Weintraub A, Widmalm G. *FEMS Microbiol Rev* 2006; **30**:382-403.
2. Croxen MA, Law RJ, Scholz R, Keeney KM, Wlodarska M, Finlay BB. *Clin Microbiol Rev* 2013; **26**:822-80.
3. Blanco JE, Blanco M, Mora A, Jansen WH, García V, Vázquez ML, Blanco J. *Vet Microbiol* 1998; **61**:229-35.
4. Penteadó AS, Ugrinovich LA, Blanco J, Blanco M, Blanco JE, Mora A, Andrade JR, Corrêa SS, Pestana de Castro AF. *Vet Microbiol* 2002; **89**:41-51.
5. Aidar-Ugrinovich L, Blanco J, Blanco M, Blanco JE, Leomil L, Dahbi G, Mora A, Onuma DL, Silveira WD, Pestana de Castro AF. *Int J Food Microbiol* 2007; **115**:297-306.
6. Peruski LF Jr, Kay BA, El-Yazeed RA, El-Etr SH, Cravioto A, Wierzbza TF, Rao M, El-Ghorab N, Shaheen H, Khalil SB, Kamal K, Wasfy MO, Svennerholm AM, Clemens JD, Savarino SJ. *J Clin Microbiol* 1999; **37**:2974-8.
7. Shashkov AS, Lipkind GM, Knirel YA, Kochetkov NK. *Magn Reson Chem* 1988;**26**:735-47.
8. Bock K, Pedersen C. *Adv Carbohydr Chem Biochem* 1983;**41**:26-66.
9. Söderman P, Jansson PE, Widmalm G. *J Chem Soc, Perkin Trans 2* 1998;639-48.
10. Jansson PE, Kenne L, Schweda E. *J Chem Soc, Perkin Trans 1* 1987;377-83.
11. Rönnols J, Pendrill R, Fontana C, Hamark C, Angles d'Ortoli T, Engström O, Ståhle J, Zaccheus MV, Säwén E, Hahn LE, Iqbal S, Widmalm G. *Carbohydr Res* 2013;**380**:156-166.
12. Iguchi A, Iyoda S, Kikuchi T, Ogura Y, Katsura K, Ohnishi M, Hayashi T, Thomson NR. *DNA Res* 2014;**22**:105-7.
13. Chad JM, Sarathy KP, Gruber TD, Addala E, Kiessling LL, Sanders DA. *Biochemistry* 2007;**46**: 6723-6732.
14. Lindquist L, Kaiser R, Reeves PR, Lindberg AA. *Eur J Biochem* 1993;**211**:763-770.
15. Marumo K, Lindqvist L, Verma N, Weintraub A, Reeves PR, Lindberg AA. *Eur J Biochem* 1992;**204**:539-545.
16. Giraud MF, McMiken HJ, Leonard GA, Messner P, Whitfield C, Naismith JH. *Acta Crystallogr D Biol Crystallogr* 1999;**55**:2043-2046.
17. Lundborg M, Modhukur V, Widmalm G. *Glycobiology* 2010;**20**:366-368.
18. Stevenson G, Neal B, Liu D, Hobbs M, Packer NH, Batley M, Redmond JW, Lindquist L, Reeves P. *J. Bacteriol.* 1994;**176**:4144-56.
19. Rojas-Macias MA, Ståhle J, Lütteke T, Widmalm G. *Glycobiology* 2015;**25**:341-347.
20. Anderson AN, Richards JC, Perry MB. *Carbohydr Res* 1992;**31**:237:249-62.

21. Marie C, Weintraub A, Widmalm G. *Eur J Biochem* 1998;**254**:378 – 381.
22. Svensson MV, Weintraub A, Widmalm G. *Carbohydr Res* 2011;**346**:449–53.

Legends to Figures

Fig. 1. ^{13}C NMR spectra (20 – 105 ppm region) of the O-polysaccharide (bottom) and the O-deacetylated polysaccharide (top). Arabic numerals refer to carbons in sugar residues denoted by letters as shown in Table 1.

Fig. 2. Part of a 2D ROESY spectrum of the O-deacetylated polysaccharide. The corresponding regions of the ^1H NMR spectrum are displayed along the axes. Arabic numerals refer to protons in sugar residues denoted by letters as shown in Table 1. Interresidue cross-peaks are annotated in bold face.

Fig. 3. Part of a 2D ^1H , ^{13}C -HMBC spectrum of the O-deacetylated polysaccharide. The corresponding regions of the ^1H and ^{13}C NMR spectra are displayed along the horizontal and vertical axes, respectively. Arabic numerals before and after oblique stroke refer to protons and carbons, respectively, in sugar residues denoted by letters as shown in Table 1. Interresidue cross-peaks are annotated in bold face.

Table 1. ^1H and ^{13}C NMR chemical shifts of the polysaccharides from *E. coli* O132 (δ , ppm)

Sugar residue	Monosaccharides						Acetyl groups	
	C-1 <i>H-1</i>	C-2 <i>H-2</i>	C-3 <i>H-3</i>	C-4 <i>H-4</i>	C-5 <i>H-5</i>	C-6 <i>H-6 (a,b)</i>	C-1	C-2 <i>H-2</i>
O-polysaccharide								
→3)-β-D-GlcpNAc-(1→ A	102.9 4.58	56.7 3.74	82.7 3.63	70.6 3.34	77.7 3.45	62.8 3.70, 3.98	175.7	23.6 2.07
→2)-α-L-Rhap-(1→ B	99.6 4.98	78.0 3.83	70.8 3.86	73.2 3.49	70.5 3.98	17.7 1.26		
→4)-α-D-Glcp-(1→ C	98.8 4.86	72.9 3.55	72.6 3.83	78.0 3.60	72.3 4.09	61.0 3.80, 3.86		
→3)-α-L-Rhap2Ac-(1→ D	98.4 4.95	72.4 5.08	77.7 3.94	71.8 3.59	70.2 4.11	17.8 1.28	174.5	21.9 2.24
→2)-α-D-Galf-(1→ E	102.3 5.18	84.7 4.07	72.7 4.35	81.4 3.86	71.4 3.76	64.0 3.61, 3.64		
O-Deacetylated polysaccharide								
→3)-β-D-GlcpNAc-(1→ A	103.1 4.72	56.7 3.88	82.4 3.70	70.0 3.50	77.7 3.51	62.3 3.76, 3.98	175.8	23.6 2.09
→2)-α-L-Rhap-(1→ B	99.6 5.01	78.0 3.86	70.8 3.87	73.2 3.50	70.5 4.00	17.7 1.27		
→4)-α-D-Glcp-(1→ C	98.8 4.88	72.9 3.58	72.8 3.84	78.4 3.62	72.3 4.07	61.3 3.75, 3.83		
→3)-α-L-Rhap-(1→ D	101.6 4.93	69.2 4.15	79.2 3.79	71.7 3.55	70.1 4.08	17.9 1.28		
→2)-α-D-Galf-(1→ E	100.8 5.23	85.6 4.16	73.2 4.38	81.7 3.88	71.9 3.79	63.9 3.62, 3.66		